AD_____

CONTRACT NUMBER: DAMD17-94-C-4082

TITLE: Increasing the Accuracy of Mammogram Interpretation

PRINCIPAL INVESTIGATOR: John A. Swets, Ph.D.

CONTRACTING ORGANIZATION: Bolt Beranek and Newman, Incorporated
Cambridge, Massachusetts 02138

REPORT DATE: October 1996

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1996	Annual 15 Sep 95 - 14 Sep 96

4. TITLE AND SUBTITLE

Increasing the Accuracy of Mammogram Interpretation

5. FUNDING NUMBERS

DAMD17-94-C-4082

6. AUTHOR(S)

Dr. John Swets

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Bolt Beranek and Newman, Inc.
Cambridge, Massachusetts 02138

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The project's aim is to refine and evaluate practically a computer-based system that will provide decision support to the radiologist in interpreting mammograms and generate a standardized report of his/her findings for the referring clinician. The fully developed system will elicit from the radiologist via spoken prompts a spoken numerical scale value for each of the diagnostically most relevant perceptual features of the mammogram (usually on a 10-point scale). It will merge the values with optimal weights via a statistical prediction rule, to calculate a probability of malignancy as an advisory for the radiologist. From the pattern of feature values, the system will automatically construct a prose report of findings for the referring clinician; spoken recommendations and impressions will be accepted by the system to complete the report. The system is to be evaluated with two groups of radiologists and cases -- one from a diagnostic or referral setting, the other from a screening or community setting -- and corresponding groups of referring clinicians. Accomplishments in Year 2 have been to set final criteria for the study cases in the referral setting, locate them in hospital files, and process them into a database; develop a much improved master list of perceptual features and associated response form; lay additional groundwork for the automated report writer; complete the software for the radiologist's interactive data entry by speech; conduct pilot image-reading sessions; and to acquire additional hardware.

14. SUBJECT TERMS	15. NUMBER OF PAGES
breast cancer, mammogram interpretation, case database, decision-support system, checklist of perceptual features, statistical prediction rule, automated report writing, data entry by speech, study case accrual, evaluation of accuracy	66
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

__X__ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

__X__ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

John A. Swets 7 Oct '96

PI - Signature Date

TABLE OF CONTENTS

INTRODUCTION

The project's aim is to refine and evaluate practically a computer-based system that will provide decision support to the radiologist in interpreting mammograms and generate a standardized report of his/her findings for the referring clinician. For each clinical case, the system provides to the radiologist a checklist of the several perceptual features of a mammogram that have been determined to be diagnostically most relevant. The system elicits from the radiologist via spoken prompts a spoken numerical scale value for each feature (usually on a 10-point scale). It will merge these values with optimal weights via a statistical prediction rule, to calculate a probability of malignancy as an advisory for the radiologist. Also, from the pattern of feature values, the planned system will automatically construct a prose report of findings for the referring clinician; spoken recommendations and impressions will be accepted by the system to complete the report.

Earlier laboratory experiments have shown that the parts of the system that aid mammogram reading and decision making -- that is, checklist and statistical prediction rule -- produce substantial increases in accuracy for the radiologist (Getty, Pickett, D'Orsi, and Swets, 1988); Swets, Getty, Pickett, D'Orsi, Seltzer, and McNeil, 1991). Clinically oriented system evaluations are now to be undertaken of refined versions of those components and, as well, of a refined and complete version of an automated-reporting capability. These evaluations will employ two groups of radiologists and cases, one representative of a community or screening setting and the other, of a referral or diagnostic setting. A significant difference between the two settings is the mix of different types of cases. Another difference is that the diagnostic center has the capability to do additional work-up (e.g., magnified views, ultrasound) while the patient is present for the mammogram, which then affect mammogram interpretation; in contrast, call-backs are required for additional work-up in the screening setting and are less frequent. It is possible that the value of the final, practical system will be viewed differently in teaching hospitals and community-oriented health maintenance organizations (HMOs).

A principal refinement of the system will be to include observed changes in perceptual features from previous to current mammograms, so-called "interval change." These changes, as to direction and amount, will themselves be feature values for the statistical prediction rule. Further, various algorithms for merging feature values -- i.e., various algorithms for defining the statistical prediction rule -- will be tested. The algorithm selected will be designed to be able to adapt over time to the mix of cases accumulating in the future in each clinical setting. Also, a speech-recognition capability is being incorporated into the system for convenient entry of case data by the radiologist. Finally, software will be developed to convert feature values into a standardized prose report of case findings; the possibility of the system's effectively recognizing concluding sentences dictated by the radiologist for the report will be explored.

Thirteen radiologists will participate as study readers, 4 in the diagnostic setting at the Brigham and Women's Hospital (BWH) and 8 in the screening setting at the Harvard Community Health Plan (HCHP). Comparison of accuracies in standard and aided conditions will be made in terms of the relative operating characteristic (ROC) (Swets, 1979; Swets and Pickett, 1982; Swets, 1988; Dorfman, Berbaum, and Metz, 1992). Focus groups of clinicians at both sites will evaluate the automatically composed reports of findings in comparison to the usual reports composed by the radiologist.

When utilized in practice, the system is expected to promote quality assurance in several respects (D'Orsi, Getty, Swets, Pickett, Seltzer, and McNeil, 1992). Its accumulating database of interpretations and outcomes will provide a detailed, quantitative, feature-by-feature basis for evaluating an individual's ability to detect and assess perceptual features that distinguish malignant and benign conditions, and for examining differences among multiple opinions. It will provide a basis for evaluating an individual's ability to set appropriate decision thresholds for different levels of treatment and for adjusting a department's average threshold, e.g., as reflected in the department's yield of biopsy recommendations (Swets, 1992). Finally, the database will allow the

tailoring of individualized tutorials for continuing education (Greenes, Swets, Getty, and Pickett, 1992).

The Body of the Report that follows contains a section on each of the main technical tasks in the project: 1. Case Selection; 2. Checklist Development; 3. Automated Report Writer; 4. Data Entry by Radiologist; and 5. Image-Reading Sessions. Remaining sections are devoted to: 6. Hardware and Software Acquisition; 7. Schedule; and 8. Personnel.

BODY OF REPORT

1. Case Selection

As outlined in the original workplan, cases have been qualified and entered into the study from two sources -- Brigham and Women's Hospital and Harvard Community Health Plan. The Human Research Committees of both institutions have approved the study protocol. Cases have been enrolled in 3 categories - malignant, benign, and "suspicious normal". The chart below summarizes eligibility criteria and enrollment statistics to date.

Category	Definition	Method of Proof	Enrolled at BWH	Enrolled at HCHP
Malignant	All types of breast cancers except lobular carcinoma-in-situ	Pathology	112	50
Benign	Focal, nonmalignant processes (i.e., benign tumors)	Pathology	156	50
Suspicious	Patient referred for additional imaging studies or accelerated follow-up and not returned to routine screening pool	Clinical/Imaging (i.e., no change in lesion appearance for 24 months of monitoring)	59	50

Patients enrolled to date fit, as expected, the demographics of the mammography referral (BWH) and screening (HCHP) practices at our two sites. This fit ensures adequate enrollment of minority groups.

1.1 Image Harvest

For each eligible case, all available original mammographic and ultrasound images at the time of the "target" examination (i.e., when the suspicious focus was identified) were harvested for use. In addition, in order to support development and evaluation of "interval change" features, mammographic images from a "comparison" examination dating approximately 12 months (range 6

to 18 months) before the target examination were also pulled. Patient-identifying information was covered by removable tape and a study number assigned to each case to ensure patient confidentiality.

1.2 Image Quality

The quality of images in each case enrolled at BWH was assessed by one of the BWH investigators (T.F.) who rated overall quality on an ordinal scale (1 to 10). In addition, this individual confirmed that all needed views were available and confirmed selection of the appropriate "comparison study."

1.3 Lesion Location

In preparation for training the statistical prediction rule, another of the BWH investigators (J.E.M.) reviewed all the selected images and listed the coordinates of the most suspicious mammographic abnormality. This step ensured that the expert readers rendered feature ratings on the same lesion.

1.4 Case Database

All available clinical mammographic and pathology data were summarized in an electronic relational database (File Maker Pro) to facilitate data extraction for the study.

1.5 Logistics of Case Collection

As we have found in prior feature-analysis studies, identifying and qualifying cases for study entry is a time- and resource-intensive task. Our entry requirements were relatively stringent and more demanding than in previous studies (including the availability of "comparison" films and, in the case of suspicious normals, "follow-up films"). Therefore, it was necessary to peruse an extremely large set of potential cases in order to find and qualify the requisite number for study entry. In fact, because of expected "leakage" from our case base (i.e., original films become

unavailable due to clinical exigencies), we qualified 10 to 20% more cases than the minimum requirement.

A further tax on resources was our requirement that every image on every eligible case be reviewed by one of the radiologist co-investigators to ensure acceptable image quality. The retrieval and presentation of all these cases proved time consuming.

2. Completion of the Master Checklist

The master checklist is the comprehensive instrument that contains all of the sections of items our participating radiologists might follow as they read cases. The particular sections followed will depend on the stage of our study and the role of the particular radiologist as either trainer or tester of our enhanced system.

2.1 Form of the Master Checklist

Work on the master checklist began in Year 1 and continued through Year 2. The finished master checklist (see Appendix) contains items for recording the status of all perceptual features considered of potential value in accurately gauging the probability that a detected lesion is malignant, as well as for reporting on that probability and the American College of Radiology's (ACR) code for recommended action. It also contains items for reporting on the presence and status of clearly benign lesions or other conditions that would need to be included in a reasonably complete radiological report back to the referring clinician.

Because descriptions of clearly benign lesions and conditions are not needed to train the statistical prediction rule, the BWH radiologists will not be required to follow those sections of the master checklist. When the HCHP radiologists read cases in our test of the effectiveness of our enhanced system, they will follow all sections, but those sections dealing with worrisome findings will have been reduced to exclude those items on the status of a suspicious lesion that were found in the training process not to be diagnostically informative.

The primary focus of work on the master checklist in Year 1 was to extend and refine its capability to support development of the computer-based statistical prediction rule. Thus, we concentrated on assembling all items of potential value in classifying suspicious lesions as benign or malignant. We considered how best to extend coverage to include interval change in these same features. We also worked carefully to translate all of our terminology into the recently approved ACR system.

2.2 Progress in Year 2

In Year 2, we extended the master checklist in several additional ways. We added items for reporting on various obviously benign lesions and other non-worrisome findings that the referring physician would want or need to hear about in addition to information on the absence or presence of worrisome findings. Also, because our enhanced system was to be applied in a real clinical and screening context, it was necessary that the questionnaire include items for reporting on, and merging into the statistical prediction rule, relevant features of the findings seen on ultrasound images.

The main remaining work in Year 2 was to review, test, and refine the questionnaire on the basis of critiques from our experts as they used the questionnaire to read actual cases. In response to the critiques, we have made numerous changes in the checklist from where it was at the end of Year 1 and even as it has evolved late into Year 2. Several changes have been made in the checklist's overall organization to better fit the sequence of considerations that experts prefer to follow, or to put particular items into sections that contain other items with which they are usually considered. For example, it seemed logical enough for the checklist designers to put the item on "Presence of Scattered Calcifications" into a section dealing with "General Breast Background," but the experts preferred that this item be moved into a restructured and more general version of a section we had narrowly devoted to "Calcification Clusters."

In addition, many items that we had inserted or planned to insert in the checklist were, in view of the burgeoning size of the checklist, dropped as being too specific or too overlapping to justify. For example, we originally aimed to have the ratings of "Percent Glandular Tissue" made separately for each breast, but soon opted for the more practical approach of getting a summary judgment for the two breasts. Similarly, we originally envisioned getting interval-change assessments on each of the key diagnostic items for every finding, but that soon appeared to be impractical. In the final version of the checklist, the radiologist reports for each worrisome finding whether it shows *any* significant change from the prior study. Only, if it is so judged, is the radiologist required to go back and assess all the key items on the prior study.

Finally -- working especially with the co-investigator who co-developed the ACR lexicon (C.J.D.) -- we made several changes aimed at increasing the clarity of the checklist items and the precision of the radiologists' reports. For example, in considering the nature of calcifications, the ACR system requires only that the radiologist check whether a particular type of calcification element is present. Initially, we followed that approach. But the critiques revealed that we could get more precise information. Now the checklist items ask for the degree to which the radiologist is confident that each type is present, and provide for the rating of confidence to be made quantitatively on a precisely labeled scale. Not only does this provide for obtaining more precise information and ultimately more accurate diagnoses in our studies, but it provides suggestions for ways that the ACR system can ultimately be improved.

We believe that we have achieved in this final version of the checklist a practical instrument of high potential informativeness for the present studies as well as an exemplary and promising system for field application.

3. Automated Report Writer

The automated report writer will render the primarily numerical information from the questionnaire in ordinary English. This translation provides a natural summary of the information

without burdening the radiologist with the task of producing a report and it can serve to standardize radiology reports so that they contain the same information and describe similar judgments from different radiologists in similar terms.

We had two major goals for Year 2 of the project:

1. Collect radiology reports and other data to determine how a radiologist would express the information in the questionnaire.

2. Design an architecture for the report writing system.

In addition to meeting both of those goals, there was a third, unintended benefit from the work, which was to contribute to refining the questionnaire. Through the text analysis of actual radiology reports we were able to notice places where the questionnaire was slightly vague, imprecise, or missing information. This led to changes in wording in several questions and the addition of a few new questions.

In this section, we describe the accomplishments in these two areas and outline our plans for the next year.

3.1 Data Analysis

We examined two kinds of data in this phase of the work. The first is targeted data, that is, reports that directly express the information from the questionnaire. The second is general data collected from radiology reports.

3.1.1 Targeted Reports

The first set of targeted data was obtained by having a radiologist fill out a questionnaire and then dictate a report of the information (for ten cases). This kind of data is the most useful because it provides direct information about how the items in the questionnaire are expressed. We are in the process of collecting a second set of targeted data that will cover the same cases, but will be

produced by a different radiologist, which will provide some information on variation in how the questions can be expressed. We will also be collecting a much larger set of reports as part of the data collection portion of the project.

Analysis of these collected reports for the text generator led us to note some improvements that could be made in the questionnaire, as shown in the following two examples. In case 181, we found that although the radiologist described the mass as "circumscribed," there was no direct item in the questionnaire that included this information. This was one factor that led to the addition of a new question. The wording of the calcification description in case 184 as "fine linear" led to a change in the wording on the questionnaire. It originally asked about "fine/linear calcifications", which implied it could be "fine *or* linear" and was changed to "fine linear" to be in keeping with the report.

Case 181: Bilateral mammograms. Dense heterogeneous glandular tissue is seen. There is a circumscribed lobular mass situated at the approximately 9 o'clock posterior position of the right breast. Remainder of the exam is within normal limits.

Impression: Mass as described above. Ultrasound should be done for further categorization. Please code as category zero (0).

Case 184: Bilateral mammograms. Dense heterogeneous glandular tissue was seen. There is a linear collection of fine linear pleomorphic calcifications in the anterior subareolar area on the left which is markedly changed from the prior exam of one year previous.

Impression: suspicious calcifications as described above. Localization biopsy recommended. Please code as category four (4).

3.1.2 General data

The general data were provided from Brigham & Women's Hospital. We have 845 reports (about 100,000 words) of these data and are using them to look at variations in how certain words

and phrases are used. We have a tool that allows us to search for all sentences that contain a certain word or phrase. For example, we found 89 sentences that have the word "circumscribed" in them. Almost all (76) say "well circumscribed," However there are some variations, such as "fairly," "smoothly," and "new" as in examples 3, 4, and 5 below. Also note that what is circumscribed is almost always described as a "nodule." However, again, there are some exceptions, as shown in 6 and 7, which are describing "masses" and "lesions."

1. The previously noted *well* **circumscribed** left upper outer quadrant nodule has remained unchanged from the examination of 12-8-88 and is consistent with a lymph node.

2. Breast ultrasound demonstrates this to be a relatively *well* **circumscribed** hypoechoic nodular density with some through transmission of sound.

3. Very posteriorly in the left slightly inner upper breast a lobulated *fairly* **circumscribed** nodule is again noted.

4. A new *smoothly* **circumscribed** nodule is seen within the right upper inner breast which was not present on the previous exam.

5. *New* **circumscribed** nodule left slightly upper outer breast which is shown to be cystic on breast ultrasound.

6. Numerous well-**circumscribed** partially obscured *masses* are present bilaterally.

7. Multiple well **circumscribed** *lesions* are compatible with cysts.

These data can be used to augment the individually collected reports in deciding how the report writer should express certain kinds of information. We can also look at how information can be combined, such as folding locations or comparisons with previous studies into a single sentence.

Although data analysis is an important part of determining how the primarily numerical information in the questionnaire can be expressed in English, we must also incorporate the guidelines provided by the American College of Radiology, because our goal is not just to make the communication natural, but to also make it consistent. The "Breast Imaging Reporting and Data System" (BI-RADS™) document provides both a lexicon and the recommended report structure that we will be incorporating into the reporting system.

3.2 Architecture

The architecture of the system should be modular, so that we can begin with a simple prototype and then be able to be incrementally extend it as more sophistication is needed. It must also be easy to change the way the system expresses information, so that in the testing phase we can have the radiologists use and critique the system, and then make adjustments.

3.2.1 Report Grammar

Our approach is to use a "grammar" to provide the top level organization of the report, as shown by the grammar rule below which is based on the BI-RADS. A grammar can be thought of a template that defines what information should be included and how it should be organized. "Non-terminals" to the left of the arrow are expanded to what is on the right of the arrow.

REPORT $\Rightarrow$ COMPOSITION FINDINGS ASSESSMENT

Each of the parts (e.g. COMPOSITION and FINDINGS) is then expanded by the grammar to include its parts, which can be words or another non-terminal in the grammar, which also must be expanded. When there are multiple choices of ways to expand a non-terminal, the choice is conditioned on the numeric values in the questionnaire. We have included a preliminary version of the questionnaire in the appendix of this document. Using that grammar, we could produce a paragraph such as the one below, which is based on Case 181.

The breast is heterogeneously dense. There is a 10 mm circumscribed lobular mass at the 9 o'clock position of the right breast. This is a new finding.

This first prototype will allow us to quickly produce simple reports so we can get early feedback. As we progress, we will incorporate a text generation system that can integrate information fluently into sentences, rather than just print out each sentence or part separately, as is done by this grammar.

3.2.2 Specialists

Another important part of our approach is isolate kinds of information that can occur in many different kinds of reports or multiple times in the same report and use text generation "specialists" to produce the descriptions. A good example of this is location. While there are many different kinds of findings, such as masses, lesions, calcifications, and architectural distortions, there are relatively few ways of describing the locations of these findings. The grammar should have one "location specialist," that can handle the various ways location is expressed, whether focal, such as o'clock position, or regional.

Determining which specialists should be created and where they are used is a complex problem, given the many different parts of the questionnaire, each with many questions. A radiologist filling in information rarely uses or even sees all of these parts, since only those for the findings he or she specifies are shown. However, the text generator must be ready to produce descriptions for all of them in case they are needed. In order to determine which specialists to create and where they should be used, we have put the many sections of the questionnaire into a table with all of the questions asked. We can then see patterns of questions that might be best handled with specialist and places where a part needs to be handled. A portion of the table is included in the appendix.

3.3 Plans

The major tasks we will be undertaking in the next year of the project are:

1. Continue collection and analysis of report data.

3. Design and implement the interface between the overall system and the report writer.

4. Implement a prototype version of the grammar, focusing on the organization of information and using canned text for the actual output.

5. Begin integrating the system with more sophisticated text generation techniques.

4. Data Entry by Radiologist

Extensive software was written to collect the test radiologists' responses to the checklist questionnaire described in Section 2 above. Data entry by the radiologist is in speech form, in response to auditory presentation of the checklist, with intelligent interactive passage through relevant portions of the checklist.

4.1 Hardware and Software

We adopted as hardware/software for speech recognition the Phonetic Engine 500 system from Speech Systems Inc. The software had bugs that we identified and communicated to the developer and we tested the new software components supplied to us to upgrade the system. The need for additional computer memory was identified and satisfied, with a resulting improvement in recognition performance. We adopted a push-to-talk method as a replacement for the system's voice-activation algorithm, with a resulting large improvement in accuracy and speed. We began testing the beta version of the package designed for Windows 95 in anticipation of moving from Windows 3.1.

4.2 Specifics of Speech Interaction

In implementing the checklist for spoken data entry, we solved the problems created by speech interaction that are not present with keyboard data entry and visual feedback. Specifically, questions and responses in written form were modified and abbreviated for the auditory format; decisions were made about how to replace keyboard input with speech input and how to provide auditory feedback in addition to visual feedback; and a process was implemented to permit the radiologist to change responses orally.

4.3 Database of Perceptual Features

A database of the perceptual features incorporated in the checklist was constructed to drive the data collection process. Written in EXCEL, the database contains information on each of the 360 questions that may be asked. It was written to permit changes to the checklist's text and its flow without requiring changes in software. It provides the question identifier; the text for screen presentation; a pointer to the digital audio file that contains the auditory prompt; a pointer to the file that contains the possible responses in auditory form; any question the present question may depend on; the next question if sequential or which of two questions to go to depending on the answer to the present question; and the name of the syntax file for the question.

4.4 Syntaxes

We developed a syntax for each of the sets of responses to the various questions, preferring to develop separate syntaxes for each question, rather than one general syntax, to assure high recognition accuracy.

4.5 Speech Recording

All the speech heard by the radiologist was recorded and digitized: including prompts to questions, feedback in the form of speaking what the recognizer recognized, and possible

responses to the questions. An index file was created to link text words used in the interactive process to the digital file containing the recorded speech.

4.6 Software for Experimental Control and Data Analysis

A program was written, using Visual Basic, to control the reading session. The program identifies the reader and the session and uses this information to open a data file specifying the cases to be read by the mammographer and their order. For each successive case, the program presents the reader with the case number and the patient's date-of-birth. It then guides the reader through the reading of the case, first identifying a lesion and then rating features associated with that lesion type, prompting the reader with a spoken cue for each feature to be rated. Feature values spoken by the reader are collected by the program and stored in a data file.

Another program was developed to read the data created in the interactive process, organize it into a full table format that represents all of the questions, and then write out the data in this format for analysis.

Constructing this collection of computer programs provided a setting in which we reviewed the checklist for consistency and logical flow.

5. Image-Reading Sessions

We scheduled image-reading sessions for the 5 participating radiologists at BWH in August, but postponed them when the case-selection and qualification processes were found to require more time and effort than planned. As a consequence, we did not collect the feature-rating data that serve to train the statistical prediction rule. We plan now to conduct the BWH image-reading sessions in the first quarter of Year 3, and to develop the merging algorithms of the statistical prediction rule and conduct HCHP baseline image-reading sessions in the second quarter.

6. Hardware and Software Acquisition

To facilitate easy viewing of current and comparison study mammograms, two mammographic viewboxes were acquired. These MAXANT Mammolume MM 0211 systems have features that are well-matched to the study requirements, including: two-tiered design (allowing film viewing in a 4-over-4 format); interior shutters (allowing viewing of 8x10 and 10x12 in film sizes simultaneously and intermittently without surrounding glare); high-intensity illumination and portability.

While no new software was acquired in Year 2, periodic upgrades to statistical and voice-recognition software were obtained.

7. Schedule and Level of Effort

The project ran at a level of effort and expenditure through Years 1 and 2 of about 90% of budget. Although parts of the project were not completed on the original schedule, we expect that they will be completed in the first few months of the next project year and that the end of Year 3 will find us on, or possibly ahead of, schedule.

8. Personnel

All staff and consultants continue with the project. Reduced time of research associate Lisa Hermann, Pharm. D., has been balanced by the addition of Mr. Edward Chao, who joined the Brigham and Women's Hospital staff after a 1994 B.S. from UCLA and two years of hospital experience.

CONCLUSIONS

Our major accomplishments in Year 2 include, first, dealing successfully with complexities and problems that arose in the definition of an appropriate set of cases for training the statistical-prediction rule at the BWH, and with the effortful requirements of selecting and qualifying individual cases to meet the definition, as well as identifying the location of the principal lesion in each.

Second, we developed a much-improved version of the checklist of perceptual features: we added items on non-worrisome findings and on ultrasound, tested the checklist on our experts, restructured it to conform to the order preferred by our experts, pruned it to consider length as well as inclusiveness, and increased its clarity.

Third, the automated report writer was advanced appreciably. Radiology reports were analyzed in relation to checklist items -- both a small sample of reports written along with a completed checklist on a case-by-case basis, and a large sample of reports written previously to look for natural semantic variation. The architecture of the report writer was designed and developed.

Fourth, extensive programming was completed to handle computer entry of spoken data by the radiologist, as the checklist is followed, and the verbal reply of the computer. A speech-recognition software package was debugged. Software was written for a database of perceptual features, for syntaxes of responses, for control of an image-reading session, and for data analysis.

Fifth, the convenience of image reading (by a dozen radiologists at two sites) was enhanced by our acquisition of two portable mammographic viewboxes, which allow intermittent reading of different-sized films without surrounding glare.

Our one regret is that the time and effort expended well on these accomplishments were so large as to prevent us from accomplishing also the formal image-reading sessions that were planned for Year 2. These sessions are now planned for the first half of Year 3.

Year 3 will contain the image-reading sessions, one of which will provide data to train a statistical-prediction rule and define the final, reduced checklist of image features, and others of which will provide baseline data and reading data enhanced by the checklist and prediction rule.

Year 3's other major task is completion and preliminary testing of the automated report writer.

Year 4 is planned to include evaluation of the report writer with focus groups of clinicians who typically receive prose reports of mammographic findings. Year 4 will also see final data analysis and preparation of the final report.

REFERENCES AND BACKGROUND LITERATURE

Dorfman, D. D., Berbaum, K. S., and Metz, C. E. (1992) Receiver operating characteristic rating analysis: Generalization to the population of readers and patients with the jackknife method. Investigative Radiology, 27, 723-731.

D'Orsi, C. J., Getty, D. J., Swets, J. A., Pickett, R. M., Seltzer, S. E., and McNeil, B. J. (1992) Reading and decision aids for improved accuracy and standardization of mammographic diagnosis. Radiology, 184, 619-622.

Getty, D. J., Pickett, R. M., Chylack, L. T., McCarthy, D. F., and Huggins, A. W. F. (1989) An enriched set of features of nuclear cataract identified by multidimensional scaling. Current Eye Research, 8(1), 1-8.

Getty, D. J., Pickett, R. M., D'Orsi, C. J., and Swets, J. A. (1988) Enhanced interpretation of diagnostic images. Investigative Radiology, 23(4), 240-252.

Getty, D. J. and Swets, J. A. (1982) Perceptual dimensions: Similarity and identification. In Geisser, H-G., Petzold, P. (Eds.) Psychophysical judgment and the process of perception. Amsterdam: North-Holland Publishing Company; and Berlin: VEB Deutscher Verlag der Wissenschaften, 1982, pp. 239-252.

Getty, D. J., Swets, J. A., Swets, J. B., and Green, D. M. (1979) On the prediction of confusion matrices from similarity judgments. Perception and Psychophysics, 26, 1-19.

Getty, D. J., Swets, J. A., and Swets, J. B. (1980) The observer's use of perceptual dimensions in signal identification. In Nickerson R. S. (Ed.) Attention and performance, VIII, pp. 361-380. Hillsdale, NJ: Lawrence Erlbaum Associates.

Getty, D. J., Swets, J. A., and Swets, J. B. (1981) Perceptual spaces revealed by multidimensional scaling procedures. In D. J. Getty, J. H. Howard, Jr. (Eds.) Auditory and Visual Pattern Recognition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Greenes, R. A., Swets, J. A., Getty, D. J., Pickett, R. M. (1992) Computer assisted instruction in radiology. Final Report submitted to Brigham and Women's Hospital. National Institute of Health, Grant No. 5 R01 CA45574-03.

Kopans, D. and D'Orsi, C. J. (1992) ACR system enhances mammography reporting. Diagnostic Imaging, September, 125-132.

Meteer, M. and McDonald, D. D. (1988) From water to wine: Generating natural language text from today's applications programs, pp. 41-49. ACL Applied Conference, Austin Texas.

Meteer, M. (1991) Bridging the 'generation gap' between text planning and linguistic realization. Computational Intelligence, 7(4), 296-304.

Meteer, M. (1991) The implications of revisions for natural language generation. In Paris, Swartuut and Mann (Eds.) Artificial Intelligence and Computational Linguistics. Klewer Academic Publishers, Boston, MA.

Meteer, M. (1991) Decision making in generation: A multi-leveled approach. Presented at the UCAI-91 Workshop on Decision Making Throughout the Generation Process, Sydney, Australia.

Meteer, M. (1992) Expressibility and the problem of efficient text planning. London: Pinter Publishers.

Meteer, M. (1994) Generating event descriptions with SAGE: A simulation and generation environment. 7th International Generation Workshop, Kennebunkport, ME.

Metz, C. E., Swets, J. A., and Al-Aish, M. (1986) Enhancing the accuracy of image-based diagnostic systems. Proposal funded under a grant from the National Cancer Institute and subcontracted from the University of Chicago. PHS Grant CA-33261.

Obuchowski, N. A. (1995) Multireader receiver operating characteristic studies: A comparison of study designs. Academic Radiology, 2, 709-716.

Seltzer, S. E., McNeil, B. J., D'Orsi, C. J., Getty, D. J., Pickett, R. M., and Swets, J. A. (1992) Combining evidence from multiple imaging modalities: A feature-analysis method. Computerized Medical Imaging and Graphics, 16(6), 373-380.

Seltzer, S. E., Getty, D. J., Tempany, C. M. C., Pickett, R, M., Schnall, M. D., McNeil, B. J., and Swets, J. A. (1995) Combined radiologist - computer system for staging prostate cancer by magnetic resonance imaging. (Submitted to New England Journal of Medicine.)

Swets, J. A. (1979) ROC analysis applied to the evaluation of medical imaging techniques. Investigative Radiology, 14(2), 109-121.

Swets, J. A. (1986a) Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. Psychological Bulletin, 99(1), 100-117.

Swets, J. A. (1986b) Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. Psychological Bulletin, 99(2), 181-198.

Swets, J. A. (1988) Measuring the accuracy of diagnostic systems. Science, 240, 1285-1293.

Swets, J. A. (1992) The science of choosing the right decision threshold in high-stakes diagnostics. American Psychologist, 47(4), 522-532.

Swets, J. A., Feehrer, C. E., Greenes, R. A., and Bynum, T. E. (1986) Use of probability estimates in medical communications and decisions. Methods of Information in Medicine, 25, 35-42.

Swets, J. A., Getty, D. J., Pickett, R. M., D'Orsi, C. J., Seltzer, S. E., and McNeil, B. J. (1991) Enhancing and evaluating diagnostic accuracy. Medical Decision Making, 11, 9-18.

Swets, J. A. and Pickett, R. M. (1982) Evaluation of diagnostic systems: Methods from signal detection theory, NY: Academic Press.

Swets, J. A., Pickett, R. M., Whitehead, S. F., Getty, D. J., Schnur, J. A., Swets, J. B., Freeman, B. A. (1979) Assessment of diagnostic technologies. Science, 205, 753-759.

Swets, J. A., Seltzer, S. E., McNeil, B. J., Tempany, C. M. C., Getty, D. J., Pickett, R. M., Schnall, M. D. (1996) Decision aids for radiologists. Paper at American Association for the Advancement of Science, Baltimore, MD, February 11.

Reader No._____

Case No._____

APPENDIX A

Response Form
X-Ray Mammography

Overview of Breast Images

Image Quality

0	1	2	3	4	5	6	7	8	9	10	OV01
poor										excellent	

Percentage of Tissue that is Glandular

_____ % OV02

Density of glandular tissue

0	1	2	3	4	5	6	7	8	9	10	OV03
slightly denser than fatty tissue										much denser than fatty tissue	

Findings?

○ **None**

○ **One or more**

If one or more findings, list in order of priority (from most suspicious to most benign) all the findings that are to be reported:

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____

Mass
MM = Mass

Special Case Densities
AT = Asymmetric Breast Tissue
TD = Tubular Density (Solitary Dilated Duct)

Architectural Distortion
AD = Architectural Distortion

Calcifications (Not-Definitely-Benign)
NC = Not-Definitely-Benign Calcifications

Calcifications (Definitely Benign)
SC = Skin Calcifications
VC = Vascular Calcifications
CC = Coarse ("Popcorn-Like") Calcifications
RL = Large Rod-Like Calcifications
LC = Spherical Lucent-Centered Calcifications
EC = Eggshell (Rim) Calcifications
MC = Milk of Calcium
UC = Suture Calcifications
DC = Dystrophic Calcifications

Secondary Signs
AA = Axillary Adenopathy
NR = Nipple Retraction
SR = Skin Retraction
ST = Skin Thickening
TT = Trabecular Thickening

Reader No._____

Case No._____

Finding No._____

Mass (MM)

● Confidence regarding the presence of a mass

0 definitely NOT present 1 2 3 4 5 6 7 8 9 10 definitely present MM01

● Distribution

- O single mass
- O multiple similar masses MM02

● Locus - breast

O left breast O right breast MM02A

● Locus - within breast (**If single mass**):

O O'clock position _____
- O anterior MM03A
- O middle MM03B
- O posterior

O central
- O anterior (subareolar)
- O middle
- O posterior

O axillary tail

● Density of mass relative to surrounding glandular tissue

0 mass density much lower 1 2 3 4 5 isodense 6 7 8 9 10 mass density much higher MM04

Mass (MM) - cont.

● Confidence about the presence of fat within the mass

0 definitely NONE present — 1 — 2 — 3 — 4 — 5 — 6 — 7 — 8 — 9 — 10 definitely some present — MM05

● Size of mass

Smallest diameter in CC view _____ mm — MM06

Smallest diameter in lateral view _____ mm — MM07

Largest diameter in CC view _____ mm — MM08

Largest diameter in lateral view _____ mm — MM09

● Shape of mass

0 round/oval — 1 — 2 — 3 — 4 — 5 lobular — 6 — 7 — 8 — 9 — 10 irregular — MM10

● Confidence that at least a small portion of the margin is indistinct due to tissue invasion

0 definitely NONE of margin indistinct due to tissue invasion — 1 — 2 — 3 — 4 — 5 — 6 — 7 — 8 — 9 — 10 definitely some of margin indistinct due to tissue invasion — MM11

● Confidence that at least a small portion of the margin is spiculated

0 definitely NOT spiculated — 1 — 2 — 3 — 4 — 5 — 6 — 7 — 8 — 9 — 10 definitely spiculated — MM12

● Percentage of the margin that is . . . (total should add to up 100%)

A. clearly circumscribed _____% — B. obscured by glandular tissue _____% — C. indistinct due to tissue invasion _____% — spiculated _____% — MM13A MM13B MM13C

● Degree of microlobulation

0 NONE — 1 — 2 — 3 — 4 — 5 — 6 — 7 — 8 — 9 — 10 extensive — MM14

Mass (MM) - cont.

● Confidence that the mass is a skin lesion

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT a skin lesion										definitely a skin lesion	MM15

● Confidence that the mass is an intramammary node

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT an intramammary node										definitely an intramammary node	MM16

● Confidence regarding the presence of related architectural distortion

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	MM17

● Confidence regarding the presence of worrisome calcifications within the mass

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	MM18

● Confidence regarding the presence of benign calcifications within the mass

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	MM19

Mass (MM) - cont.

- Is there an ultrasound?
  - O Yes
  - O No

MMUL

Questions relating to Ultrasound

- Appearance of the mass wall (ultrasound)
  - O well circumscribed
  - O indistinct
  - O irregular

MMUL1

- Contents of the mass (ultrasound)
  - O solid
  - O indeterminate
  - O cystic

MMUL2

- Response of the posterior wall of the mass (ultrasound)
  - O enhancement
  - O iso-echoic
  - O shadowing

MMUL3

- Shape of the mass (ultrasound)
  - O round
  - O ellipsoid
  - O irregular

MMUL4

Mass (MM) - cont.

Relationship to Prior Study

- This mass finding is: MM20 MM21
  - O new
  - O significantly changed
  - O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Asymmetric Breast Tissue (AT)

● Confidence regarding presence of asymmetric breast tissue

0	1	2	3	4	5	6	7	8	9	10	AT01
definitely NOT present										definitely present	

● Distribution of asymmetric breast tissue

O focal AT02

O global

● Locus - breast

O left breast O right breast AT03

● Locus - within breast (**If focal distribution**): AT04A AT04B

O O'clock position _____

- O anterior
- O middle
- O posterior

O central

- O anterior (subareolar)
- O middle
- O posterior

O axillary tail

Asymmetric Breast Tissue (AT) - cont.

- Size of distribution of asymmetric breast tissue (**If focal distribution**)

  Largest diameter in CC view _____ mm AT05

  Largest diameter in lateral view _____ mm AT06

- Confidence regarding the presence of worrisome calcifications within the asymmetric breast tissue

0 definitely NOT present	1	2	3	4	5	6	7	8	9	10 definitely present

AT07

- Confidence regarding the presence of benign calcifications within the asymmetric breast tissue

0 definitely NOT present	1	2	3	4	5	6	7	8	9	10 definitely present

AT08

- Confidence regarding presence of related architectural distortion

0 definitely NOT present	1	2	3	4	5	6	7	8	9	10 definitely present

AT09

Relationship to Prior Study

- This asymmetric breast tissue finding is: AT10 AT11

  - O new
  - O significantly changed
  - O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Tubular Density (Solitary Dilated Duct) (TD)

● Confidence regarding the presence of a tubular density (solitary dilated duct)

0 definitely NOT present	1	2	3	4	5	6	7	8	9	10 definitely present

TD01

● Locus - breast

O left breast O right breast TD02

● Locus - within breast

O O'clock position _____ O anterior O middle O posterior TD03A TD03B

O central O anterior (subareolar) O middle O posterior

O axillary tail

Relationship to Prior Study

● This finding is: TD04 TD05

O new

O significantly changed

O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Architectural Distortion (AD)

● Confidence regarding presence of architectural distortion in a least one view

0	1	2	3	4	5	6	7	8	9	10	AD01
definitely NONE present										definitely some present	

● Locus - breast

O left breast O right breast AD02

● Visible in both views or only one?

O both

O one: Oblique

O one: CC

AD03

● Locus - within breast

If both views: O O'clock position _____ AD04A

O anterior
O middle
O posterior AD04B

O central
- O anterior (subareolar)
- O middle
- O posterior

O axillary tail

If Oblique: AD05A

O superior
O middle
O inferior
O axillary tail

If CC: AD05B

O medial
O middle
O lateral

Architectural Distortion (AD) - cont.

● Confidence that the architectural distortion is related to prior surgery

0 1 2 3 4 5 6 7 8 9 10 AD06
definitely NOT related to prior surgery — definitely related to prior surgery

● Confidence regarding the presence of related worrisome calcifications

0 1 2 3 4 5 6 7 8 9 10 AD07
definitely NOT present — definitely present

● Confidence regarding the presence of related benign calcifications

0 1 2 3 4 5 6 7 8 9 10 AD08
definitely NOT present — definitely present

● Confidence regarding the presence of a related mass

0 1 2 3 4 5 6 7 8 9 10 AD09
definitely NOT present — definitely present

● Confidence regarding the presence of related asymmetric breast tissue

0 1 2 3 4 5 6 7 8 9 10 AD10
definitely NOT present — definitely present

Relationship to Prior Study

● This architectural distortion finding is: AD11 AD12

- O new
- O significantly changed
- O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Calcifications (Not-Definitely-Benign) (NC)

● Confidence regarding the presence of not-definitely-benign calcifications

0	1	2	3	4	5	6	7	8	9	10	NC01
definitely NOT present										definitely present	

● Distribution of calcifications (choose one):

- O single focal distribution (clustered, linear, segmental)
- O multiple similar focal distributions
- O regional distribution
- O diffuse NC02

● Locus - breast

O left breast O right breast NC03

● Locus - within breast (**If single focal or regional distribution**):

NC04A
NC04B
NC04C
NC04D

If single focal

O O'clock position _____

- O anterior
- O middle
- O posterior

O central

- O anterior (subareolar)
- O middle
- O posterior

O axillary tail

If regional

(check as many as apply)

- ☐ superior medial
- ☐ superior lateral
- ☐ inferior medial
- ☐ inferior lateral
- ☐ central
  - ☐ anterior (subareolar)
  - ☐ middle
  - ☐ posterior
- ☐ axillary tail

Calcifications (Not-Definitely-Benign) (NC)-cont.

Elemental Characteristics

● Size of largest individual element (best visual estimate) NC05

less than 0.5 mm | 0.5 mm to 1.0 mm | greater than 1.0 mm

● Variability of size of elements

0 1 2 3 4 5 6 7 8 9 10 NC06

0: low variability of size

10: high variability of size

● Degree to which the calcifications can be characterized as fine linear

0 1 2 3 4 5 6 7 8 9 10 NC07

0: definitely NONE of the elements are fine linear

9 10: at least one or two elements definitely are, or several probably are, fine linear

● Degree to which the calcifications can be characterized as branching

0 1 2 3 4 5 6 7 8 9 10 NC08

0: definitely NONE of the elements are branching

9 10: at least one or two elements definitely are, or several probably are, branching

● Degree to which the calcifications can be characterized as pleomorphic (heterogeneous)

0 1 2 3 4 5 6 7 8 9 10 NC09

0: definitely NONE of the elements are pleomorphic (heterogeneous)

8 9 10: at least one or two elements definitely are, or several probably are, pleomorphic (heterogeneous)

● Degree to which the calcifications can be characterized as amorphous (indistinct)

0 1 2 3 4 5 6 7 8 9 10 NC10

0: definitely NONE of the elements are amorphous (indistinct)

8 9 10: at least one or two elements definitely are, or several probably are, amorphous (indistinct)

Calcifications (Not-Definitely-Benign)(NC) - cont.

Elemental Characteristics (continued)

- Degree to which the calcifications can be characterized as punctate

0 1 2 3 4 5 6 7 8 9 10 NC11

0: definitely NONE of the elements are punctate

10: at least one or two elements definitely are, or several probably are, punctate

- Degree to which the calcifications can be characterized as round

0 1 2 3 4 5 6 7 8 9 10 NC12

0: definitely NONE of the elements are round

10: at least one or two elements definitely are, or several probably are, round

Distributional Characteristics (**If focal or multiple similar focal distributions**)

- Number of elements

less than 5 5 to 10 more than 10 NC13

- Size of the focal distribution

Largest dimension in CC view _____ mm NC14

Largest dimension in lateral view _____ mm NC15

- Degree to which the distribution can be characterized as clustered

0 1 2 3 4 5 6 7 8 9 10 NC16

0: definitely NOT clustered

10: definitely clustered

- Degree to which the distribution can be characterized as linear

0 1 2 3 4 5 6 7 8 9 10 NC17

0: definitely NOT linear

10: definitely linear

Calcifications (Not-Definitely-Benign)(NC) - cont.

Distributional Characteristics **(If focal or multiple similar focal distributions)** (continued)

- Degree to which the distribution can be characterized as segmental

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT segmental										definitely segmental	NC18

Relationship to Other Aspects of This Study **(If focal, multiple similar focal, regional distribution)**

- Confidence regarding presence of related architectural distortion

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	NC19

- Confidence regarding presence of related mass or asymmetric breast tissue

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	NC20

Relationship to Prior Study

- This not-definitely-benign calcifications finding is: NC21 NC22
  - O new
  - O significantly changed
  - O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Skin Calcifications (SC)

● Confidence regarding the presence of skin calcifications

0	1	2	3	4	5	6	7	8	9	10	
definitely not present	NONE									definitely present	SC01

● Locus - breast

○ left breast ○ right breast SC02

● Locus - quadrant (check as many as apply)

☐ superior medial

☐ superior lateral

☐ inferior medial

☐ inferior lateral SC03

Relationship to Prior Study

● This skin calcifications finding is: SC04 SC05

○ new

○ significantly changed

○ unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Vascular Calcifications (VC)

● Confidence regarding the presence of vascular calcifications

0	1	2	3	4	5	6	7	8	9	10	VC01
definitely NOT present										definitely present	

● Locus - breast

O left breast O right breast VC02

● Locus - within breast

O O'clock position _____ O anterior VC03A
O middle VC03B
O posterior

O central O anterior (subareolar)
O middle
O posterior

O axillary tail

Relationship to Prior Study

● This vascular calcifications finding is: VC04 VC05

O new

O significantly changed

O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Coarse ("Popcorn-Like") Calcifications (CC)

● Confidence regarding the presence of coarse ("popcorn-like") calcifications

0	1	2	3	4	5	6	7	8	9	10	CC01
definitely NOT present										definitely present	

● Distribution of calcifications:

- O single focal distribution (clustered, linear, segmental)
- O multiple similar focal distributions
- O regional distribution
- O diffuse CC02

● If single focal distribution or multiple similar focal distributions, the shape of the distribution(s) is/are best described as:

- O clustered
- O linear
- O segmental CC03

● Locus - breast

O left breast O right breast CC04

Coarse or "Popcorn-Like" Calcifications (CC) - cont.

● Locus - within breast **(If focal or regional distribution)**

CC05A
CC05B
CC05C
CC05D

If focal

- O O'clock position _____
  - O anterior
  - O middle
  - O posterior
- O central
  - O anterior (subareolar)
  - O middle
  - O posterior
- O axillary tail

If regional
(check as many as apply)

- ☐ superior medial
- ☐ superior lateral
- ☐ inferior medial
- ☐ inferior lateral
- ☐ central
  - ☐ anterior (subareolar)
  - ☐ middle
  - ☐ posterior
- ☐ axillary tail

Relationship to Prior Study

● This coarse ("popcorn-like") calcifications finding is: CC06 CC07

- O new
- O significantly changed
- O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Large Rod-Like Calcifications (RL)

● Confidence regarding the presence of large rod-like calcifications

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	RL01

● Distribution of calcifications (choose one):

- O single focal distribution (clustered, linear, segmental)
- O multiple similar focal distributions
- O regional distribution
- O diffuse RL02

● If single focal distribution or multiple similar focal distributions, the shape of the distribution(s) is/are best described as:

- O clustered
- O linear
- O segmental RL03

● Locus - breast

O left breast O right breast RL04

Large Rod-Like Calcifications (RL) - cont.

● Locus - within breast (**If single focal or regional distribution**):

RL05A
RL05B
RL05C
RL05D

If single focal	**If regional (check as many as apply)**
O O'clock position _____	☐ superior medial
O anterior	☐ superior lateral
O middle	☐ inferior medial
O posterior	☐ inferior lateral
O central	☐ central
O anterior (subareolar)	☐ anterior (subareolar)
O middle	☐ middle
O posterior	☐ posterior
O axillary tail	☐ axillary tail

● Number of elements

less than 5 | 5 to 10 | more than 10 RL06

Relationship to Prior Study

● This large rod-like calcifications finding is: RL07
RL08

O new

O significantly changed

O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Spherical Lucent-Centered Calcifications (LC)

● Confidence regarding the presence of spherical lucent-centered calcifications

0	1	2	3	4	5	6	7	8	9	10	LC01
definitely NOT present										definitely present	

● Distribution of calcifications (choose one):

- O single focal distribution (clustered, linear, segmental)
- O multiple similar focal distributions
- O regional distribution
- O diffuse LC02

● If single focal distribution or multiple similar focal distributions, the shape of the distribution(s) is/are best described as:

- O clustered
- O linear
- O segmental LC03

● Locus - breast

O left breast O right breast LC04

Spherical Lucent-Centered Calcifications (LC) - cont.

● Locus - within breast **(If single focal or regional distribution)**

LC05A
LC05B
LC05C
LC05D

If single focal	**If regional (check as many as apply)**
O O'clock position _____	☐ superior medial
O anterior	☐ superior lateral
O middle	☐ inferior medial
O posterior	☐ inferior lateral
O central	☐ central
O anterior (subareolar)	☐ anterior (subareolar)
O middle	☐ middle
O posterior	☐ posterior
O axillary tail	☐ axillary tail

● Size of the spherical lucent-centered calcifications

Largest dimension in CC view _____ mm LC06

Largest dimension in lateral view _____ mm LC07

● Number of elements

less than 5 5 to 10 more than 10 LC08

Relationship to Prior Study

● This spherical lucent-centered calcifications finding is: LC09
LC10

O new

O significantly changed

O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Eggshell (Rim) Calcifications (EC)

- Confidence regarding the presence of eggshell rim calcifications

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	EC01

- Distribution of calcifications (choose one):

  - O single focal distribution (clustered, linear, segmental)
  - O multiple similar focal distributions
  - O regional distribution
  - O diffuse EC02

- If single focal distribution or multiple similar focal distributions, the shape of the distribution(s) is/are best described as:

  - O clustered
  - O linear
  - O segmental EC03

- Locus - breast

  O left breast O right breast EC04

Eggshell (Rim) Calcifications (EC) - cont.

- Locus - within breast **(If focal or regional distribution)**

EC05A
EC05B
EC05C
EC05D

If focal

- O O'clock position _____
  - O anterior
  - O middle
  - O posterior
- O central
  - O anterior (subareolar)
  - O middle
  - O posterior
- O axillary tail

If regional
(check as many as apply)

- ☐ superior medial
- ☐ superior lateral
- ☐ inferior medial
- ☐ inferior lateral
- ☐ central
  - ☐ anterior (subareolar)
  - ☐ middle
  - ☐ posterior
- ☐ axillary tail

- Number of elements

less than 5 | 5 to 10 | more than 10 — EC06

Relationship to Prior Study

- This eggshell (rim) calcifications finding is: — EC07 EC08

  - O new
  - O significantly changed
  - O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Milk of Calcium (MC)

● Confidence regarding the presence of milk of calcium

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	MC01

● Distribution of calcifications (choose one): MC02

- O single focal distribution (clustered, linear, segmental)
- O multiple similar focal distributions
- O regional distribution
- O diffuse MC02

● If single focal distribution or multiple similar focal distributions, the shape of the distribution(s) is/are best described as:

- O clustered
- O linear
- O segmental MC03

● Locus - breast

O left breast O right breast MC04

Milk of Calcium (MC) - cont.

- Locus - within breast (**If focal or regional distribution**):

MC05A
MC05B
MC05C
MC05D

If focal	**If regional (check as many as apply)**
O O'clock position _____	☐ superior medial
O anterior	☐ superior lateral
O middle	☐ inferior medial
O posterior	☐ inferior lateral
O central	☐ central
O anterior (subareolar)	☐ anterior (subareolar)
O middle	☐ middle
O posterior	☐ posterior
O axillary tail	☐ axillary tail

- Number of elements

less than 5 | 5 to 10 | more than 10 MC06

Relationship to Prior Study

- This milk of calcium finding is: MC07 MC08

O new

O significantly changed

O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Suture Calcifications (UC)

● Confidence regarding the presence of suture calcifications

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	UC01

● Locus - breast UC02

○ left breast ○ right breast

● Locus - quadrant (check as many as apply)

☐ superior medial

☐ superior lateral

☐ inferior medial

☐ inferior lateral UC03

Relationship to Prior Study

● This suture calcifications finding is: UC04 UC05

○ new

○ significantly changed

○ unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Dystrophic Calcifications (DC)

- Confidence regarding the presence of dystrophic calcifications

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT present										definitely present	DC01

- Distribution of calcifications:

  - O single focal distribution (clustered, linear, segmental)
  - O multiple similar focal distributions
  - O regional distribution
  - O diffuse DC02

- If single focal distribution or multiple similar focal distributions, the shape of the distribution(s) is/are best described as:

  - O clustered
  - O linear
  - O segmental DC03

- Locus - breast

  O left breast O right breast DC04

Dystrophic Calcifications (DC) - cont.

● Locus - within breast (**If focal or regional distribution**):

DC05A
DC05B
DC05C
DC05D

If focal

- ○ O'clock position _____
  - ○ anterior
  - ○ middle
  - ○ posterior
- ○ central
  - ○ anterior (subareolar)
  - ○ middle
  - ○ posterior
- ○ axillary tail

If regional
(check as many as apply)

- ☐ superior medial
- ☐ superior lateral
- ☐ inferior medial
- ☐ inferior lateral
- ☐ central
  - ☐ anterior (subareolar)
  - ☐ middle
  - ☐ posterior
- ☐ axillary tail

● Number of elements

less than 5 | 5 to 10 | more than 10 DC06

Relationship to Prior Study

● This dystrophic calcifications finding is: DC07
DC08

- ○ new
- ○ significantly changed
- ○ unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Skin Thickening (ST)

● Confidence regarding presence of skin thickening

0	1	2	3	4	5	6	7	8	9	10	
definitely not present	NONE									definitely some present	ST01

● Locus - breast

O left breast O right breast ST02

● Visible in both views or only one?

O both
O one: Oblique
O one: CC ST03

● Locus - quadrant (check as many as apply)

If both views: ☐ superior medial ST04A
☐ superior lateral
☐ inferior medial
☐ inferior lateral

If Oblique: ☐ superior ST04B
☐ inferior

If CC: ☐ medial ST04C
☐ lateral

Relationship to Prior Study

● This skin thickening finding is: ST05 ST06

O new
O significantly changed
O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Skin Retraction (SR)

● Confidence regarding the presence of skin retraction

0	1	2	3	4	5	6	7	8	9	10	SR01
definitely NONE present										definitely some present	

● Locus - breast

O left breast O right breast SR02

● Visible in both views or only one?

O both
O one: Oblique
O one: CC SR03

● Locus - quadrant (check as many as apply)

If both views: ☐ superior medial SR04A
☐ superior lateral
☐ inferior medial
☐ inferior lateral

If Oblique: ☐ superior SR04B
☐ inferior

If CC: ☐ medial SR04C
☐ lateral

Relationship to Prior Study

● This skin retraction calcifications finding is: SR05 SR06

O new
O significantly changed
O unchanged or not significantly changed

Case No._____

Finding No._____

Nipple Retraction (NR)

● Confidence regarding the presence of nipple retraction

0	1	2	3	4	5	6	7	8	9	10	NR01
definitely NONE present										definitely some present	

● Locus - breast

O left breast O right breast NR02

Relationship to Prior Study

● This nipple retraction finding is: NR03 NR04

O new

O significantly changed

O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Axillary Adenopathy (AA)

● Confidence regarding presence of axillary adenopathy

0	1	2	3	4	5	6	7	8	9	10	AA01
definitely NONE present										definitely some present	

● Locus - breast

O left breast O right breast AA02

Relationship to Prior Study

● This axillary adenopathy finding is: AA03

AA04

O new

O significantly changed

O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

Trabecular Thickening (TT)

● Confidence regarding presence of trabecular thickening

0	1	2	3	4	5	6	7	8	9	10	TT01
definitely NONE present										definitely some present	

● Locus - breast

O left breast O right breast TT02

Relationship to Prior Study

● This trabecular thickening finding is: TT03

TT04

O new

O significantly changed

O unchanged or not significantly changed

Reader No._____

Case No._____

Finding No._____

ACR CATEGORY

Please select the appropriate ACR category:

- O I (negative exam, routine follow-up)
- O II (benign finding, routine follow-up)
- O III (probably benign finding, accelerated follow-up)
- O IV (suspicious finding, biopsy recommended)
- O V (finding highly suggestive of malignancy, biopsy recommended)

OVERALL DIAGNOSTIC JUDGMENT

Benign vs. Malignant---Rate the likelihood (as the number of chances in 100) that the finding is indicative of malignancy:

Rating (0 to 100) _____

where: 0 = certainly benign or normal

100 = certainly malignant

APPENDIX B

Mammography Report Grammar

The following is a portion of the grammar that is used to select and organize information in the reports. Non-terminals to the left of the arrow are expanded to what is on the right of the arrow. When there are choices, those choices are marked with conditions, which use the numerical values in the questionnaire to select the correct option.

Top Level Organization

REPORT ⇒ COMPOSITION FINDINGS ASSESSMENT

Report Body

COMPOSITION ⇒ S1 :condition (OV02 0-1) |

S2 : condition (OV02 2-4) |

S3 : condition (OV02 5-8) |

S4 : condition (OV02 9-10)

FINDINGS ⇒ FINDING | FINDING FINDINGS | S5

FINDING ⇒ MASS | ;conditioned on OV03

CALCIFICATION-NB |

CALCIFICATION-B |

ARCHITECTURAL-DISTORTION

Mass

MASS ⇒ MASS-DESCRIPTION ASSOCIATED-FINDINGS

MASS-DESCRIPTION ⇒ "There is a" SIZE TYPE+MODIFIER "mass" LOCATION.

SIZE ⇒ VALUE :condition (MM06) "cm"

TYPE+MODIFIER ⇒ "circumscribed" | "lobular" | "glandular" ;conditioned on values

Calcifications (Not definitely benign)

CALCIFICATION-NB ⇒ CALC-NB-DESCRIPTION

Location Specialist

LOCATION ⇒ "at" O'CLOCK :condition (MM03 xxx) "in the" LOCUS "breast" |
"in the " CENTRAL :condition (MM03 xxx) "position of the" LOCUS "breast" |
"in the" AXILLARY :condition (MM03 xxx) "in the" LOCUS "breast"

O'CLOCK ⇒ VALUE :condition (MM03 xxx) "o'clock"

CENTRAL ⇒ "anterior" :condition (MM03 xxx) |
"middle" :condition (MM03 xxx) |
"posterior" :condition (MM03 xxx)

AXILLARY ⇒ "axillary tail"

LOCUS ⇒ "right" :condition (MM01 R) |
"left" :condition (MM01 L) |

Overall Assessment

ASSESSMENT ⇒ IMAGE-QUALITY REL-TO-PRIOR

Tissue Densitiy

S1 ⇒ "The breast is almost entirely fat."

S2 ⇒ "There are scattered fibroglandular densities."

S3 ⇒ "The breast is heterogeneously dense."

S4 ⇒ "The breast is extremely dense."

No Findings

S5 ⇒ "No masses, significant calcifications, or other abnormalities are visible."

Relationship to Prior Studies

S6 ⇒ "This is a new finding."

S7 ⇒ "The finding is significantly changed."

S8 ⇒ "The finding is unchanged."

Questionnaire Analysis

The following table shows each question in the total questionnaire in the rows and the different specific subparts of the questionnaire in the columns. The item number is filled in when a question is asked in that subpart. This allows us to see patterns and isolate specialists in the text generator when the same information is elicited in many different parts of the form.

QUESTION	MM	AT	TD	AD	NC	SC	VC	CC	RL	LL	EC	ST	SR	RD	AA	TT
confidence regarding presence	01	01	01	01	01	01	01	01	01	01	01	01	01	01	01	01
Locus- breast	02	03	02	02	03	02	02	04	04	04	04	02	02	02	02	02
Locus within breast	03	04	03	03	04		03	05	05	05	05					
Locus-quadrant						03						03	03			
Distribution of calc.					02			02	02	02	02					
Distr of calcs								03	03	03	03					
Distributional: num. elts					13				06	07	06					
Density wR\rt surrounding	03															
Confidence about fat	04															
Size	6-9	5-6			05											
shape	10															
Confidence spiculaetd	12															
degree of microlobulatoin	13															
Confidence skin lesion	14															
Confidence intrammary	15															
Confidence related distortion	16	09														
Condifence worrisome calcs	17	07		05	09											
Confidence benign calc	18	08		06												
Relationship to prior study	19	10	04	09	21	04	04	06	07	08	07	04	04	03	03	03
Distribution of breast tissue		02														
Conf related to surgery				04												
Presence related mass				07												
Pr. rleates assynmetric tissue				08	20											
Variability of size					06											
Degree fine linear					07											
Degree branching					08											
Degree pleomorphic					09											
Degree amorphous					10											
Degree punctate					11											
Degree round					12											
size of focal collec					14-15					5-6						
Degree collection clustered					16											
Degree collecton linear					17											
Degree collecton segmental					18											

PROJECT STAFF

John A. Swets (P.I.), PhD	BBN Corporation
David J. Getty, (Co P.I.) PhD	BBN Corporation
Ronald M. Pickett, PhD	BBN Corporation
Marie Mateer PhD	BBN Corporation
Barbara Freeman, BA	BBN Corporation
Steven E. Seltzer, (Co P.I.) MD	Brigham and Women's Hospital
Lisa Hermann, Pharm D	Brigham and Women's Hospital
Edward Chao, BS	Brigham and Women's Hospital
William J. Otto, Jr., MD	Harvard Community Health Plan
Carl J. D O'rsi, MD	University of Massachusetts Medical Center